

DOCUMENT RESUME

ED 261 807

PS 015 404

AUTHOR Miller, Michael D.; Ottinger, Donald R.
 TITLE Influence of Labeling on Ratings of Infants: A Prematurity Prejudice.
 PUB DATE Aug 85
 NOTE 15p.; Paper presented at the Annual Meeting of the American Psychological Association (Los Angeles, CA, August 23-27, 1985).
 PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
 EDRS PRICE MF01/PC01 Plus Postage.
 DESCRIPTORS Behavior Rating Scales; *Bias; Child Caregivers; *Infants; *Labeling (of Persons); *Premature Infants; Stereotypes; *Student Attitudes; *Undergraduate Students; Videotape Recordings

ABSTRACT

Two full term and two preterm infants were videotaped while being administered six items from the Brazelton Scale. Infants were assigned alternately the labels "preterm" and "fullterm" and shown to a group of 256 undergraduate students. It was hypothesized that: (1) subjects who view infants labeled as preterm would rate them lower on objective scales of behavioral performance and on subjective scales of infant characteristics; (2) subjects who rate the infant's performance on subjective scales would demonstrate a greater labeling bias than subjects who rate the infant's performance on the objective, behaviorally anchored Brazelton Scale; and (3) subjects who view a labeled preterm would report less confidence in their ability to care for that baby's needs compared to subjects viewing a labeled fullterm infant. Results reveal that subjects demonstrated no influence of label in their ratings of the infants' Brazelton performance or in assessing their confidence in performing various caregiving tasks for the particular infant viewed. However, on global ratings, the labeled preterms were rated as significantly less attentive, less healthy, smaller, more difficult to care for, less enjoyable to interact with, and less sociable. Results are discussed as supporting the existence of a prematurity prejudice. (Author/RH)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED261807

PS 015404

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Do Not Copy or Cite without written permission from the authors.

Influence of Labeling
on Ratings of Infants:
A Prematurity Prejudice¹

Michael D. Miller*

Donald R. Ottinger

Purdue University

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Michael D.
Miller

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Paper presented at the annual meeting of
The American Psychological Association,
Los Angeles, California, August, 1985

*Please submit all correspondence to Michael D. Miller, Ph.D.,
Department of Pediatrics, Arnett Clinic, 2600 Greenbush Street,
Lafayette, IN 47904

Abstract

Two fullterm and two preterm infants were videotaped while being administered six items from the Brazelton Scale. Infants were assigned alternately the labels "preterm" and "fullterm" and shown to a group of 256 undergraduate students. Subjects demonstrated no influence of label in their ratings of the infants' Brazelton performance, or in assessment of their confidence in performing various caregiving tasks for the particular infant viewed. However, on global ratings, the labeled preterms were rated as significantly less attentive, less healthy, smaller, more difficult to care for, less enjoyable to interact with, and less sociable. Results are discussed as supporting the existence of a prematurity prejudice.

Influence of Labeling on Ratings of
Infant Behavior: A Prematurity Prejudice

Accompanying the rise in reported preterm infant viability rates (Kopp and Parmelee, 1979) have come reports of delayed or arrested developmental advances in this population (Holmes, Nagy, Slaymaker, Sosnowski, Prinz, & Pasternak, 1982; Field, Ting, & Shuman, 1979; Goldstein, Caputo, & Taub, 1976). Preterm infants have been observed to be more fussy at time of discharge (Friedman, Jacobs, Werthman, 1982) and are reported by their mothers to be less socially interactive at date of hospital discharge and more difficult to console at age 4 months and 8 months (Field, Hallock, Dempsey, Dabiri, & Shuman, 1978; Field, Hallock, Dempsey, & Shuman, 1978). There have been reports of the face and cry features of preterm infants being evaluated by parents as more irritating, annoying, and disturbing - a finding the authors interpreted as indicating preterms to be perceived as more aversive than fullterms (Frodi, Lamb, Leavitt, Donovan, Neff, & Sherry, 1978).

In an attempt to probe for a "prematurity stereotype," Stern and Hildebrandt (1984) found that when mothers and other adults were told that an infant had been premature, the infant was rated as weaker, slower, more shy, less competent, and less liked than fullterms. However, many of the items queried (cuddliness, friendliness, assertiveness) bore little relation to the behavior viewed (playing alone with a single toy). Subjects therefore did not need to attend to the behaviors viewed in order to make their judgments; hence the observed behavior may have been superfluous to the ratings. Subjects may have responded, instead, based on their conjecture of how most preterm or fullterm infants ought to be perceived. While the study does offer evidence of a prematurity stereotype (i.e., pre-set conventional notion of a group), it cannot be said solely on the basis of this study that parents would assess their preterm infant's behavior on the basis of their preterm birth status alone, regardless of the infant's behavior.

To ignore, or to color one's perception of the behavior itself, and to act instead upon the expectations set by a prematurity stereotype would constitute a prematurity prejudice -- an action based on an attitude developed or maintained despite objective evidence to the contrary (Allport, 1954; Ehrlich, 1973).

In light of these limitations, the present research attempted to investigate whether adults discriminate against preterm infants by allowing their judgment of an infant's behavior to be heavily influenced by knowledge of a label status independent of, or in spite of, actual emitted behavior. In the present study this possibility of a preterm prejudice was investigated at two levels: One, subjects were asked to evaluate an infant's elicited performance on the objective tasks of the Brazelton Neonatal Behavioral Assessment Scale. Two, subjects were asked also for

their more global ratings of the infant's overall characteristics. The present study differs from the Stern and Hildebrandt study by showing younger infants interacting with an adult in a standardized format. This study also differs in that it features both fullterm and preterm infants.

It was hypothesized that:

1. subjects who view infants labeled as preterm would rate them lower on objective scales of behavioral performance and on subjective scales of infant characteristics,
2. subjects who rate the infant's performance on the subjective scales would demonstrate a greater labeling bias than subjects who rate the infant's performance on the objective, behaviorally anchored Brazelton Scale, and
3. subjects who view a labeled preterm would report less confidence in their ability to care for that baby's needs compared to subjects viewing a labeled fullterm infant.

Method

Subjects

Two hundred and fifty-six college students (128 males, 128 females) served as subjects in return for partial course credit. Subjects ranged in age from 18 to 29 years ($\bar{X}$ = 19.5) and 96.3% were single. Less than 2% of the sample were parents at the time of the study, but 96% reported that they plan to have children in the future.

Stimulus Materials

Six separate color videotapes were used, four of which each depicted one infant to be rated, two of which were instructional tapes. The Brazelton Instructional Tape consisted of an examiner administering six items² of the Brazelton Neonatal Behavioral Assessment Scale (BNBAS) to a healthy, white fullterm infant 9 days old. Prior to each of the six items, an experimenter appeared on the screen to introduce the upcoming item and to alert the subjects for which behaviors to look for in order to help them understand how the item was scored. While the examiner was shown administering the items to the infant, a voice-over accented scorable features of the infant's behavior as they occurred.

The Likert Instructional Tape featured the same experimenter introducing each item and explaining in general terms the performance objective for the infant. This tape differed from the

first in that it provided no details of specific performance to watch for, nor did it feature an infant being administered the items. No infant was featured on this tape so as to preclude subjects from using the demonstration infant as a comparison or norm. This was intended to enhance the subjectivity of the rating process for this group.

The other four videotapes comprised the Rating Procedure Tapes (data collection phase). Each depicted one healthy, white infant (9-11 days corrected age) being administered the BNEAS in the family home. The same female examiner administered the six items in the exact order to all five (1 training and 4 rating) infants. Two infants were the product of fullterm births; two of the infants were born prematurely (34 and 37 weeks gestational age). Infants were depicted for a total of approximately 11-13 minutes each.

Procedure

During the procedure, all subjects were seated at tables approximately 4-6 meters from a 24-inch color videotape monitor. Subjects were run in groups of 8 or less and were visually isolated from each other by partitions. Subjects were randomly assigned to the Brazelton Rating Group or to the Likert Rating Group.

Brazelton Rating Group. These 128 subjects first viewed the Brazelton Instructional Tape. Subsequent to the instructional tape and prior to the rating procedure, subjects were read one of two scenarios introducing the infant they were about to rate--either a healthy, fullterm infant or a healthy, preterm infant.³ Both oral scenarios introduced the baby as Chris (Kris) and the infant's sex was never identified. The scenarios, or labels, were systematically varied to an equal number of subjects who then viewed a videotape of one of the four infants (also systematically varied). Each infant was presented to an equal number (8) of male and female subjects in the Likert and Brazelton groups under both label conditions. Therefore each subject rated only one infant under one labeling condition in this between-subjects design.

After each of the six items, subjects in this group were allowed up to one minute to score the item on an adapted (some language was simplified) Brazelton scale utilizing the Scale's behavioral descriptors. Specific descriptors accompanying each scoring level were arranged in a 9-point Likert-scale format with "Best Possible," "Average," and "Worst Possible" performance identified. At the completion of the sixth and final item, subjects completed questionnaires requesting information on their experience with infants. They were asked also how confident they would feel in performing a number of caregiving tasks for the individual infant viewed, and they were instructed to rate that infant on a number of subjective indices. Thus, for example, subjects in the preterm group were asked to describe that particular baby, not preterm

babies in general.

Likert Rating Group. These 128 subjects were not given an instruction period. Instead they were told that the experimenters were interested in their opinion of how well the baby performed. Subjects were then read one of the two scenarios described above to label the infant viewed. Preceding each of the six behaviors, these subjects viewed a short segment of the Likert Instructional Tape which briefly introduced the upcoming item. After each item, subjects were allowed 30 seconds to rate the infant's performance on 9-point Likert scales. Upon completion of the sixth item, these subjects completed the same questionnaire as the Brazelton group. In addition, they rated the viewed infant on Stern and Hildebrandt's (1984) series of global 7-point Likert scales asking for their general impressions of the infant's characteristics. Again, subjects were instructed to rate the infant just viewed (Table 1 lists the individual items).

Results

Results will be reported in four categories: Manipulation Check, Ratings on the six Brazelton items, Confidence Levels, and Subjective Ratings.

Manipulation Check

The label was manipulated successfully. More than 95% of the subjects recalled the infant's labeled birth status correctly, 96% were able to report the infant's name (Chris or Kris) correctly, but less than one-half correctly identified the sex of the infant rated.

Brazelton Ratings

A 2 (label) x 2 (rating instrument) x 2 (sex of subject) analysis of variance was conducted on the data. ANOVA's conducted on the six items individually or in total revealed no significant main effect for label among the Brazelton subjects, the Likert subjects, or the two groups combined ($p > .10$). Sex of subject also exerted little influence on the data except on item 6 (Consolability with Intervention) where females rated all infants significantly higher than males $F(1,248) = 5.74$, $p < .05$. The rating tool used did exert a significant influence on five of the items ($p < .01$), excepting item 5 (Pull to Sit). However, in breaking down the sample into Likert and Brazelton group subjects, neither group's data revealed a significant labeling effect. That is, some of the items were scored higher by the Brazelton subjects and others were scored higher by the Likert subjects. However these differences were independent of the labeling condition.

To assess individual infant influences, in the next ANOVA a

fourth factor, Infants, was added. On all six items and on the total, the individual infants were found to have a significant main effect ($p < .001$) on subjects' ratings. The two preterm infants received similar scores under both labeling conditions. By contrast, the two fullterm infants' scores were more disparate, with one receiving lower scores under the fullterm label, unlike any of the other three infants.

Confidence Labels

After rating the infant's Brazelton scale performance, each subject was asked how confident he or she would feel in performing each of thirteen separate caregiving tasks for the particular infant viewed. These items ranged from specific tasks such as taking this baby's temperature to more general tasks such as being a good parent to this baby.

Analysis of variance revealed that the label affixed to the infant viewed was not a significant factor in the subjects' ratings of any of the thirteen tasks individually or in sum ($p > .10$). On twelve of the items, female subjects reported significantly higher confidence levels than did male subjects. There was, however, a significant interaction effect between label and sex of subject for the sum total of the confidence ratings. Males reported a lower level of confidence in caring for labeled preterms than they did for labeled fullterms, while females tended to report similar levels of confidence for both groups $F(1,248) = 3.89$, $p < .05$. Subjects' reported levels of experience in eight areas of infant caregiving were significantly related ($r = .57$, $p < .001$) with their total scores on the confidence scale. However, neither reported confidence or experience demonstrated any relation to ratings on the Brazelton items ($r = .01$ and $.08$, respectively, $p > .10$).

Insert Table 1 about here

Subjective Ratings

On the questionnaire, each subject was asked to rate the one infant viewed on three subjective variables. Analyses of variance revealed significant main effects for label on all three variables. Infants labeled as preterm were rated as less fun/more annoying to interact with $F(1,242) = 4.23$, $p < .05$, more difficult to care for $F(1,242) = 21.23$, $p < .001$, and less healthy $F(1,242) = 28.27$, $p < .001$ than labeled fullterms.

Analyses of variance were conducted next on the sixteen global 7-point rating scales completed by the Likert subjects (only). Infants labeled preterm were rated as significantly smaller, less attentive, less friendly/more shy, less cute, and less assertive ($p < .05$). Labeled preterms tended to be rated as somewhat weaker, more fussy, and less fun to play with ($p < .10$). For the total of the sixteen items, labeled preterms received significantly less optimal ratings than did the labeled fullterms ($p < .05$). Stern and Hildebrandt (1984) have combined these sixteen items into four clusters (see Table 1). Labeled preterms received significantly less optimal scores on the four Physical Characteristics items $F(1,124) = 7.45$, $p < .01$, on the four Sociability items $F(1,124) = 5.49$, $p < .05$, and on the three Cognitive Competency items $F(1,124) = 6.76$, $p < .01$. Ratings on the fourth cluster, Behavioral Characteristics, were not significantly influenced by label $F(1,124) = 1.38$, $p > .10$. Sex of subject was a significant factor on all four clusters ($p < .05$), as females accorded higher ratings than male subjects to all infants.

Discussion

The first hypothesis, that subjects would rate labeled preterms lower on all rating scales, received partial support. These infants were rated more poorly on subjective scales, but not on ratings of behavioral performance (Brazelton items). Accordingly, the second and third hypotheses were not supported. The rating instrument used had no effect on a labeling bias and subjects reported no less confidence in their ability to care for a preterm infant relative to a fullterm infant. Labeling did have a significant effect on subjects' ratings of the infant's health, temperament, physical characteristics, sociability, and attentiveness. Thus, in spite of the fact that the subjects had previously reported no behavioral differences (compared to other subjects) they nonetheless proceeded to rate labeled preterms more poorly on subjective indices.

Although the labeled preterm infants were rated as no different than the labeled fullterms as regards their Brazelton performance, the labeled preterms were nonetheless adjudged to be smaller, less healthy, less attentive, less enjoyable to interact with, and less sociable than the labeled fullterms. Though labeled preterms were expected to be more difficult to care for, subjects did not express any less confidence in their ability to do so.

These results, in support of Stern and Hildebrandt's (1984) findings, demonstrate that there is a bias against preterms. However, this bias does not universally color all perceptions. In addition to the Brazelton items and the confidence levels, label had no effect upon subjects' ratings of such qualities as happiness, alertness, or cuddliness. That the bias was not observed on all dependent variables adds credence to the strength of the bias

observed on the other measures. It is possible to obtain a strong artifactual labeling effect by presenting subjects with minimal or no information other than a label (e.g., Learning Disabled, Schizophrenic, Mentally Retarded) (Aloia, Maxwell, & Aloia, 1981; Critchley, 1979).

Rather than demonstrate true prejudice toward these individuals, such findings only demonstrate that when there is no other evidence upon which to base an evaluation, subjects will expect a certain level of performance from a child based on the label that the child has presumably earned. The present failure to find bias on the Brazelton items suggests that the duration of exposure to the infant's behavior was sufficient to rate the behavior on its own merits. It also indicated that the bias observed on the other items cannot be explained away as a mere artifact of experimental contrivance. Subjects were accorded sufficient opportunity to observe and consider adequately the quality of the performance.

It was suggested above that the subjects in Stern and Hildebrandt's study may have been conjecturing about the general class of preterms (stereotyping) rather than demonstrating actual prejudice in the face of contradictory data. This possibility arises from the fact that many of their rating items were not subject to support or contradiction by the depicted behaviors. The present study attempted to determine whether that stereotype would so permeate subjects' judgment that perception and/or reporting of specific, observable performance would be so altered. It did not. The problem in interpretation, then, is why the stereotypic view remained firm enough to support the other biases after the behavior had already been observed and reported to be of average or better-than-average quality. Stereotypic views not reversible by contradicting evidence denotes prejudice (Allport, 1954).

If the findings of this study can be extrapolated to the natural parent-infant relationship, then the tenacity of these biases in the face of observed and reported evidence to the contrary is somewhat disturbing. Possibly these stereotypic views are not abandoned even after actual observation of the infant's competent behavior. It would be in this manner then, that the interactions of preterm infants and their mothers could continue to be different from that of fullterm mother-infant dyads long after initial contact. This would allow for the development of self-fulfilling prophecies over time (Green & Solnit, 1964).

The question of why some items generated bias while others did not is open to speculation. Perhaps the label manipulation was too weak to override the salience of the observed behavior. The Brazelton rating task required close attention to the stimuli; rating criteria were specific and objective. Judgments were to be made on the basis of what was seen, rather than on the basis of what

one supposes to be the case, as in the subjective ratings. These subjective ratings were global impressions, not all necessarily tied to the behavior viewed. Possibly such evaluative judgments are more specific behavior.

For whatever reason(s), those stereotypic views did remain after the behavioral rating experience and apparently influenced the subjects' ratings of their overall impressions of the infant viewed. That this bias remained after the non-biased Brazelton ratings does lend support to a determination of preterm birth prejudice -- attitudes held despite objective evidence to the contrary.

Interpretation of these results must be made with caution given that they are based on the self-reports of non-parents in an artificial situation. Nonetheless, the social interactive properties of the Brazelton scale items and the expressed parentage plans of the subjects indicate that this analogue study at least approximates the naturally occurring interaction situation of parent and infant.

These results, in conjunction with Stern and Hildebrandt's (1984) findings, do offer evidence that a prematurity bias is operable in the judgments of some adults and could be rather resilient to passive reception of disconfirming data. As such,

References

- Allport, G.W. (1954). The nature of prejudice. Garden City, NY: Addison Wesley.
- Aloia, G.F., Maxwell, J.A., & Aloia, S.D. (1981). Influence of a child's race and the EMR label on initial impressions of regular classroom teachers. American Journal of Mental Deficiency, 85, 619-623.
- Brazelton, T.B. (1973). Neonatal behavioral assessment scale. Philadelphia: J. B. Lippincott.
- Critchley, D.L. (1979). The adverse influence of psychiatric diagnosis labels on the observation of child behavior. American Journal of Orthopsychiatry, 49, 157-160.
- Field, T.M., Hallock, N., Dempsey, J., Dabiri, C., & Shuman, H.H. (1978). A first year follow-up of high risk infants. Child Development, 49, 119-131.
- Field, T.M., Hallock, N., Dempsey, J., & Shuman, H.H. (1978). Mothers' assessments of term infants and preterm infants with Respiratory Distress Syndrome. Child Psychiatry and Human Development, 9, 75-83.
- Field, T.M., Ting, G., & Shuman, H.H. (1979). The onset of rhythmic activities in normal and high risk infants. Developmental Psychobiology, 12, 97-100.
- Friedman, S.L., Jacobs, B.S., & Werthman, M.W. (1982). Preterms of low medical risk. Infant Behavior and Development, 5, 3-10.
- Frodi, A.M., Lamb, M.E., Leavitt, L.A., Donovan, W.L., Neff, C. & Sherry, D. (1978). Fathers' and mothers' responses to the faces and cries of normal and premature infants. Developmental Psychology, 14, 490-498.
- Goldstein, K.M., Caputo, D.V., & Taub, H.B. (1976). The effects of prenatal and perinatal complications on development at one year of age. Child Development, 47, 613-621.

Green, M., & Solnit, A.J. (1964). Reactions to the threatened loss of a child: A vulnerable child syndrome. Pediatrics, 34, 58-66.

Holmes, D.L., Nagy, J.N., Slaymaker, F., Sosnowski, R.J., Prinz, S.M., & Pasternak, J.F. (1982). Early influences of prematurity, illness, and prolonged hospitalization on infant behavior. Developmental Psychology, 18, 744-750.

Kopp, C.B. & Parmelee, A.H. (1979). Prenatal and perinatal influences on infant behavior. In J.D. Osofsky (Ed.), Handbook of infant development, (pp. 29-75). New York: Wiley & Sons.

Stern, M., & Hildebrandt, K.A. (1984). Prematurity stereotype: Effects of labeling on adults' perceptions of infants. Developmental Psychology, 20, 360-362.

Footnotes

¹This manuscript is based on a thesis submitted by the first author to the Department of Psychological Sciences, Purdue University, in partial fulfillment of the requirements for the doctoral degree. The authors are grateful to Debra Bendell and Rebecca Smith for their review of an earlier version of this manuscript.

²Response Decrement to Light, Orientation Animate--Visual, Orientation Animate--Auditory and Visual, Defensive Movements, Pull-to-Sit, Consolability with Intervention.

³This is Chris, a (fullterm baby who was born right on time) (baby who was born six weeks premature): Proper prenatal care, healthy pregnancy, two healthy parents. Chris had a normal birth for a (fullterm baby) (premature baby and spent some time in a newborn intensive care unit until ready to go home). Chris was tested as a part of routine hospital procedure shortly after being discharged. The test was administered within one week (after birth) (after Chris was scheduled to be born, about seven weeks after the actual birth).

Table 1

Subjective ratings of infants labeled as "Preterm"

Questionnaire item	p
Less enjoyable to interact with	.05*
More difficult to care for	.001***
Less healthy	.001***
Smaller	.001***
Weaker	.10
Large/Fine Featured	n.s.
Less Cute	.05*
PHYSICAL CHARACTERISTICS (less optimal)	.01**
Less Friendly	.01**
Sad/Happy	n.s.
Less Assertive	.01**
Less Fun to Play with	.10
SOCIABILITY (less)	.05*
Dumb/Smart	n.s.
Less Attentive	.001***
Slow/Fast Learner	n.s.
COGNITIVE COMPETENCY (lower)	.01**
Loud/Quiet	n.s.
More Fussy	.10
Sleepy/Alert	n.s.
Poorly/Well-Behaved	n.s.
Cuddly/Not Cuddly	n.s.
BEHAVIORAL CHARACTERISTICS	n.s.